

Docket No.: YOR920030230US1

INTERNATIONAL BUSINESS MACHINES CORPORATION

APPLICATION FOR UNITED STATES LETTERS PATENT

entitled

**PERFORMANCE PREDICTION SYSTEM WITH QUERY
MINING**

by

**David M. Chess
Sophia Krasikov**

PERFORMANCE PREDICTION SYSTEM WITH QUERY MINING

TECHNICAL FIELD OF THE INVENTION:

- 5 This invention relates to a system, method and service that collects and stores information regarding queries received by a performance prediction system, and that enables the stored information to be queried to provide enhanced performance prediction results to a user.

10 BACKGROUND OF THE INVENTION:

- Reputation and performance prediction systems are known and increasingly relied upon for providing valuable information. Examples include the Better Business Bureau, and Internet applications such as online auction and other types of user rating
15 systems. Typically, these systems are intended to provide users with information about the reputation of businesses and/or the likely success of a transaction. To use these systems, a user typically requests pertinent information. In online embodiments, requests are typically submitted by sending queries of various kinds to the system, and receiving appropriate responses.

20

If any information relating to the queries is retained, it is most likely used only in a limited fashion, such as for internal record-keeping and load-analysis purposes.

- What is needed is an improved performance prediction system that makes use of the
25 received queries for providing additional and enhanced information to interested parties.

SUMMARY OF THE INVENTION:

- 30 The foregoing and other problems are overcome by methods and apparatus in accordance with embodiments of this invention.

The inventors have recognized that valuable information can be learned from the queries received by a performance prediction system. The information can prove useful to businesses and other interested parties for gaining insight into the needs, interests and/or perspectives of system users. Information on trends and assorted other information may be produced as well from the query record. Examples of useful information include assessing the popularity of a product or service, and assessing the concerns of a typical customer.

Disclosed herein are methods and apparatus for a performance prediction system and service. A typical user of the system disclosed herein may be a business, such as a manufacturer of goods for retail distribution. However, the class of potential users is not restricted to business entities, as other types of organizations and individuals may benefit from the use of this invention.

Aspects of the teachings herein include a system for producing enhanced performance prediction information. The system includes a source of data that contains performance prediction data, where the data source is coupled to a query component adapted for execution of at least one query. Also provided is a data gathering component adapted for collecting query-relevant data concerning received queries and for storing the query-relevant data in a submitted query database. Also included is a meta-query component that is coupled to the submitted query database and that is adapted for querying the query-relevant data and for producing enhanced performance prediction information.

In a further aspect of this invention, a performance prediction service uses acquired knowledge of the queries that are received to make predictions concerning the future performance of certain entities, such as business entities. By analyzing query patterns and the actual or predicted performance of business entities, the performance prediction service is enabled to observe correlations between queries, query patterns and performance that may be used to facilitate the prediction of future performance.

Another aspect of the teachings herein includes a method for producing enhanced performance prediction information, by: querying a source of performance prediction data; collecting query-relevant data about each query; storing the query-relevant data; and, querying the query-relevant data for producing the enhanced performance prediction information.

This invention also provides a computer program product stored on a computer readable media, for implementing a method to receive queries of performance prediction information, store query-relevant data derived from the queries, receive further queries directed to the query-relevant data, and provide the query-relevant data, and/or information derived from the query-relevant data, as enhanced performance prediction information.

In a further aspect, this invention provides a service that is accessible over a data communications network, such as the Internet. The service is operable for receiving queries from users that are related to the performance of some entity, for storing query-relevant data derived from the queries, for receiving queries that are related to the stored query-relevant data, and for providing results for the queries of query-relevant data as enhanced performance prediction information to the users over the data communications network, or via some other means.

Another aspect of this invention includes a communications device for receiving enhanced performance prediction information from a service, where the device includes a communications interface adapted for at least one of submitting a meta-query to the service and receiving the enhanced performance prediction information in response to a meta-query. In the preferred embodiment the service includes a source of data that includes performance prediction data, the data source being coupled to a query component adapted for receiving the meta-query and executing the meta-query; a data gathering component adapted for collecting query-relevant data for fulfilling meta-queries, and storing the query-relevant data in a submitted query database; and a meta-query component coupled to the submitted query database adapted for querying the query-relevant data to produce the enhanced performance

prediction information in response to the reception of the meta-query from the communications device.

The various embodiments of the invention described herein are illustrative only, and are
5 not to be considered as limiting the scope or application of this invention.

BRIEF DESCRIPTION OF THE DRAWINGS:

The above set forth and other features of the invention are made more apparent in the
10 ensuing Detailed Description of the Invention when read in conjunction with the attached Drawings, wherein:

Fig. 1 is a block diagram that depicts components of a performance prediction service;

15 **Fig. 2** is a flow chart illustrating operation of the performance prediction service for providing a customer with performance prediction data; and

Fig. 3 is a flow chart illustrating operation of the performance prediction service for storing performance prediction data.

20

DETAILED DESCRIPTION OF THE INVENTION

Disclosed herein are methods and apparatus for an enhanced performance prediction data service. An “enhanced performance prediction data service,” as used herein, refers
25 to a service for providing a user, such as an individual or an organization, with information concerning queries received by the performance prediction data service. An organization may be a for-profit or a not-for-profit organization, including government organizations. The performance prediction data service may be used by third parties, or it may be a “captive” service that is operated and used by a single individual or
30 organization.

“Performance prediction data” generally refers to data relevant to the operation of businesses and other organizations that is typically maintained in a database and made available to users through operation of one or more queries. Performance prediction data is so named as it is typically useful as an aid in predicting the reliability, performance, or some such aspect of a business or other organization. For the purposes of this invention “enhanced performance prediction data” generally refers to information that includes at least one component part that is derived from information descriptive of received queries for performance prediction data. Enhanced performance prediction data is so named as it is typically useful for providing insight not available in the performance prediction data, or for serving as an enhancement to the performance prediction data. Both performance prediction data and enhanced performance prediction data are generated and returned to a user in response to queries, which are referred to herein, for convenience only, as one of a “performance query” and a “meta-query.”

In general terms, a performance query is directed to obtaining the performance prediction data, while a meta-query is directed to obtaining the enhanced performance prediction data. The performance prediction data is generated by querying a data source, such as a database of product information. The enhanced performance prediction information is derived from a stored historical record of previously received queries. The foregoing definitions are not to be construed as being limitations on the practice of this invention, and are provided simply for the sake of clarity and convenience.

Fig. 1 illustrates one embodiment of a data processing system for implementing a performance prediction system 101. In this embodiment, the system 101 communicates with a communications device 105 through a network 104, such as the global Internet. Although illustrated as a PC terminal, the communications device 105 may include any type of device, including both mobile and fixed communications devices, capable of generating queries and/or meta-queries, and of receiving query and meta-query results from the system 101. The communications device 105 is assumed to include a suitably programmed communications interface 105A adapted for at least one of submitting a query, such as a meta-query, to the performance prediction

system 101 and for receiving performance prediction and enhanced performance prediction information from the performance prediction system 101.

During operation of the system 101, a user of the communications device 105
5 submits, via communications interface 105A, a query 120 to the system 101, seeking one of performance prediction information 125 or enhanced performance prediction information 135. In this embodiment, where the user is seeking performance prediction information 125, the query 120 is referred to as a performance query (PQ) 121. If the user is seeking enhanced performance prediction information 135, the
10 query 120 is referred to as a meta-query (MQ) 122. The system 101 determines the type of query 120, and directs processing of the query 120 accordingly.

The system 101 further includes a query component 110, a data gathering component 112 and a performance prediction component 113. The data gathering component 112
15 is bidirectionally coupled to the query component 110 and to a submitted query database 102, containing submitted query-related data 132. The performance prediction component 113 is bidirectionally coupled to the query component 110, to the submitted query database 102, and to a database of performance prediction information 103, containing performance prediction data 133.

20
In the instance where the query 120 is the performance query 121, the query is routed to the query component 110 for execution. The query component 110 operates with the data gathering component 112 and with the performance prediction component 113. The data gathering component 112 gathers query-relevant information 132 about
25 submitted performance queries 121, and records the query-relevant information 132 in the submitted query database 102. The performance prediction component 113 obtains data 132, 133 from the submitted query database 102 and the database of performance prediction information 103, respectively, as necessary to satisfy the performance query 121. The performance prediction component 113 may employ any
30 suitable techniques for analyzing the data 132, 133, such as statistical techniques, to obtain query patterns and correlations between query patterns. The performance prediction component 113 derives at least one of an actual or predicted performance

of an entity of interest (such as a business or an individual), and may consider the correlations between queries, query patterns and performance to facilitate the prediction of future performance of the entity of interest. Results are returned to the query component 110 for routing to the communications device 105 as performance prediction information 125.

The meta-query component 111 provides for the submission of a second type of query 120, the meta-query 122, to request the enhanced performance prediction information 135. The enhanced performance prediction information 135 is produced by querying query-relevant data 132 stored in the submitted query database 102. Query-relevant data 132 may include a variety of data types, such as copies and/or numbers of individual received queries 120, aggregations of query-related information (e.g., times and locations from which queries are received), and also the results of statistical analyses of queries 120. The query-relevant data 132 may further include external information 130, such as, in non-limiting examples, temporal information describing the date or time of query reception, and user location(s) and user-related identifying information (e.g., TCP/IP address). Combinations of such data may be provided as is determined suitable. Results are returned to the meta-query component 111 for routing to the communications device 105 as the enhanced performance prediction data 135.

The performance prediction information 125 and the enhanced performance prediction information 135 are typically routed to the communication device 105 submitting the query 120, however, this is not required and the results may be directed elsewhere.

As an example of performance prediction data 133 and query-relevant data 132, one embodiment involves the non-limiting context of a manufacturer and/or distributor of retail goods. In this context, the user may request performance prediction information 125 such as the price of a certain item from a plurality of suppliers, an identification of competitive items, information related to durability and reliability, information related to complaints received from other users about the goods and/or the supplier of

the goods, and other similar types of information. Another user may consider it important to obtain enhanced performance prediction information 135, in order to gain insight into market trends in the interests of consumers, features that are of importance to consumers, the perception of the supplier in the eyes of consumers, and other similar types of information. The enhanced performance prediction information preferably includes an information component that is derived from the submitted query database 102 that contains the submitted query-related data 132.

Fig. 2 is a logical flow diagram illustrating an aspect of the operation of the performance prediction system 101. When a query 120 is communicated to the system 101, it is received in step 201. The system 101 examines the query 120 and determines at step 202 the type of query. That is, whether the query 120 is a performance query 121 or a meta-query 122. If the received query 120 is a meta-query 122, then the query is sent to the meta-query component 111 for execution, otherwise it is sent to the query component 110. At step 203 the query component 110 makes use of the submitted query database 102 and the database of performance prediction information 103, as necessary to satisfy the performance query 121. At step 204 the meta-query component 111 executes the meta-query 122 by reference to the submitted query database 102.

In the preferred embodiment information about the received query 120 is also recorded at step 205 in the submitted query database 102, regardless of the type of query 120, and query results are returned at step 206. It should be recognized that at least some of the operations shown in **Fig. 2** may be rearranged, such as by executing step 205 prior to step 202, or after step 206.

Fig. 3 is a logical flow diagram illustrating operation of a presently preferred embodiment of the system 101. In step 301, a performance query 121 is received by the query component 110. The performance query 121 is satisfied in a two step process. First, the database of performance relevant information 103 is interrogated at step 302, then the submitted query database 102 is interrogated at step 303. At step 304 the performance query 121 is recorded in the submitted query database 102. The

result that is returned to the user in step 305 is, at least potentially, the result of considering information derived from previous queries 120, which has been recorded in the submitted query database 102. As was described with respect to **Fig. 2**, some of the steps may be rearranged, such as by executing step 304 prior to step 302, or after
5 step 205, or by executing step 303 prior to step 302.

Another example of performance prediction data 133 and query-relevant data 132 involves the non-limiting context of a rating system. For example, in some on-line systems buyers and sellers can purchase goods or services, or can exchange goods or
10 services, and can also rate various aspects of transactions. For example a potential buyer sends a performance query 121 to ascertain performance prediction information related to a potential seller. The returned information may include information related to the seller's reputation in previous transactions, customer satisfaction survey information, and other such data. The buyer or the seller using the system 101 may
15 issue a meta-query 122 to ascertain enhanced performance prediction information. The enhanced information includes, for example, information descriptive of queries 120 received from previous potential buyers concerning the seller, queries received that included only the seller in the results, queries received that excluded the seller from the results, the number of queries received about the seller, copies of received
20 queries, copies of results returned in response to received queries, how many queries were received during some prescribed period of time; patterns of queries received (e.g., queries received from specific locations, or from specific users), and correlations between queries.

25 The system 101 supports techniques for producing enhanced performance prediction information 135 that includes at least one of: filtering query-relevant data 132, aggregation of query-relevant data 132, statistical analysis of query-relevant data 132, and a search of query-relevant data 132 for locating information stored for a single query 121 (or for multiple queries 120). Accordingly, the examples herein are merely
30 exemplary and not limiting of reporting results.

Additional examples include embodiments where the results 125, 135 are filtered, or edited, so as to remove sensitive information, such as personal identification information. Filtering may occur at any stage, as appropriate. For example, the filtering may occur by foregoing the storage of certain data, by generally excluding an ability to query for certain data, or by excluding certain data when generating the query report. In other embodiments, the information 125, 135 is not filtered.

The system 101 may include a discovery function 141 that enables users to obtain a list of companies that meet certain criteria.

In a preferred embodiment the system 101 determines the identity and/or address of the user sending the query 120, and establishes whether that user is authorized to access the system 101, and/or arranges to charge the user for the service provided.

Further examples of producing enhanced performance prediction information 135 include analysis of query-relevant data 132 to produce, as examples only, information that is descriptive or indicative of: the number of received queries that are about a specific user (possibly during some prescribed period of time); about a specific product; queries that include or exclude an item of interest (such as an identity of a business or a competitor of the business); queries that are indicative of trends; a most often asked question received in queries; queries that are about reliability (product and/or business); queries related to estimations of a success or failure of a transaction; and an average, a minimum, a maximum, and/or a count of queries containing a single (or multiple) specific subject(s).

In additional embodiments of this invention, the meta-query component 111 allows users to register to be notified when certain information 132 comes to exist in the submitted query database 102. For instance, a user might register to receive notification whenever some predetermined criteria has been reached, such as when a some number of queries 120 about the user have been recorded in the database 102, or when any query is received about the user, or when some number of queries concerning a competitor have been recorded. A business user may also be interested

in being informed whenever a query is received about a certain product that indicates that the buyer of the product may require service or repair of the product (e.g., a query is received from buyer requesting information about the availability of service or parts suppliers.

5

In another embodiment, a user might register to automatically initiate a meta-query 122 at a specified time. In this case the meta-query may be already stored in the system 101, and is activated at the appropriate time.

10 As can be appreciated, in one aspect this invention enables the performance prediction system 101 to examine and analyze the submitted query database 102 that contains the submitted query-related data 132. This can be seen to provide an ability to perform “data mining” on the submitted query-related data 132, or more generally to permit a “query mining” operation to be performed.

15

One skilled in the art should recognize that the invention is not limited to a performance prediction data system 101, as disclosed herein, and that the teachings herein may be employed in a variety of embodiments. For example, a business may integrate the teachings herein with internal information systems, and therefore not deal with a service provider 101. The teachings herein may be used as a part of a predictive maintenance information system, where maintenance planning or an engineering staff monitors access to information by maintenance personnel. The teachings herein may be implemented without a network 104 in a standalone system. Accordingly, it is considered that these and other additional embodiments are within the teachings of this invention, as defined

20

by the appended claims.

25